

Operationalising AI ethics: barriers, enablers and next steps

Year: 2021 | Citations: 150 | Authors: J. Morley, Libby Kinsey, Anat Elhalal, Francesca Garcia, M. Ziosi

Abstract

By mid-2019 there were more than 80 AI ethics guides available in the public domain. Despite this, 2020 saw numerous news stories break related to ethically questionable uses of AI. In part, this is because AI ethics theory remains highly abstract, and of limited practical applicability to those actually responsible for designing algorithms and AI systems. Our previous research sought to start closing this gap between the ‘what’ and the ‘how’ of AI ethics through the creation of a searchable typology of tools and methods designed to translate between the five most common AI ethics principles and implementable design practices. Whilst a useful starting point, that research rested on the assumption that all AI practitioners are aware of the ethical implications of AI, understand their importance, and are actively seeking to respond to them. In reality, it is unclear whether this is the case. It is this limitation that we seek to overcome here by conducting a mixed-methods qualitative analysis to answer the following four questions: what do AI practitioners understand about the need to translate ethical principles into practice? What motivates AI practitioners to embed ethical principles into design practices? What barriers do AI practitioners face when attempting to translate ethical principles into practice? And finally, what assistance do AI practitioners want and need when translating ethical principles into practice?